

Język SQL dla analityków danych w środowisku PostgreSQL

Cele szkolenia

Na szkoleniu uczestnik zdobędzie wiedzę na temat funkcjonowania mechanizmów dostępnych w środowisku PostgreSQL odpowiedzialnych za wyszukiwanie i analizowanie danych. Uczestnik posiędzie wiedzę odnośnie tworzenia konstrukcji wybierających dane z wielu tabel, wykorzystywania wbudowanych funkcji jednowierszowych i grupowych, uwzględnienia formatowania danych oraz ich sortowania. Wiedza dotyczy także operacji zbiorowych i podzapytań, jak również wyrażeń tablicowych oraz stosujących wielowymiarowe grupowania. Wiedza i umiejętności zdobywane podczas kursu obejmują również wykorzystanie funkcji analitycznych do zaawansowanego przetwarzania danych i ich analizy.

Umiejętności

Dzięki szkoleniu uczestnik będzie potrafił:

- Wykonywać złożone zapytania w celu wyświetlenia danych z pojedynczych tabel oraz na podstawie różnych wariantów złączeń,
- Realizować raporty oparte o podzapytania proste i skorelowane,
- Wykonywać złożone operacje z wykorzystaniem wspólnych wyrażeń tablicowych,
- Tworzyć raporty z wielopoziomowymi agregacjami, za pomocą klauzul Rollup, Cube i Grouping Sets,
- Zapamiętywać wyniki operacji SQL w tabelach oraz zapisywać złożone polecenia w postaci widoków,
- Wykorzystywać dane z plików zewnętrznych w działaniach SQL oraz zapisywać wyniki operacji SQL w plikach CSV.

Profil uczestników

Kurs przeznaczony jest dla słuchaczy, którzy przeprowadzają operacje analityczne lub tworzą raporty w środowisku PostgreSQL. Szkolenie jest kierowane do wszystkich, którzy potrzebują przeprowadzać analizę danych i podejmować decyzje biznesowe.

Przygotowanie uczestników

Od uczestników szkolenia wymagana jest znajomość środowiska Windows.

Szczegółowy program szkolenia

Język SQL - wybieranie danych

Tworzenie zapytań

- Słowa kluczowe
- Składnia poleceń
- Stosowanie aliasów dla kolumn i tabel

Sortowanie danych

Użycie funkcji w zapytaniach

- Funkcje jednowierszowe
- Funkcje grupowe
- Konwersja typów danych

Wyrażenie CASE i jego zastosowania

Wybieranie danych z wielu tabel

- Sposoby łączenia tabel
- Rodzaje złączeń
- Równościowe (equi-join)
- Nierównościowe (non-equi-join)
- Typy złączeń zewnętrznych (outer-join): lewostronne, prawostronne, obustronne
- Samozłączenie (self-join)

Operacje zbiorowe na wynikach zapytań

- Część wspólna (Union)
- Przecięcie (Intersect)
- Różnica (Except)

Podzapytania

- Lokalizacja podzapytań
- Proste
- Skorelowane
- Rola podzapytań skorelowanych w złożonych zapytaniach

Funkcje analityczne

- Zastosowanie
- Partycje
- Okna
- Funkcje rankingowe
- Funkcje okna
- Funkcje LAG/LEAD
- Funkcje FIRST_VALUE, LAST_VALUE i NTH_VALUE
- Funkcja NTILE
- Funkcja WIDTH_BUCKET

Wyrażenia tablicowe w złożonych zapytaniach

- Idea
- Klazula WITH
- Przykłady zastosowań

Wielowymiarowe grupowanie

- Idea i zastosowania
- Rollup
- Cube
- Grouping Sets

Przechowywanie danych wynikowych

- Widoki i ich rola w dostępie do danych
- Tworzenie tabel trwałych
- Tworzenie i zastosowanie tabel tymczasowych

Współpraca z danymi zewnętrznymi

- Import danych w plików zewnętrznych CSV
- Eksport wyników zapytań do plików CSV

Metoda realizacji szkolenia

Szkolenie realizowane jest w formie naprzemiennie następujących po sobie mini wykładów oraz ćwiczeń praktycznych. Szkolenie łączy w sobie fachową wiedzę merytoryczną z praktycznymi przykładami jej wykorzystania w środowisku pracy.

Liczba dni, liczba godzin szkoleniowych

3 dni, 24 godziny szkoleniowe

Ścieżka rozwoju po szkoleniu

Po zakończeniu szkolenia rekomendowane jest skorzystanie z:

- Szkolenie: „PostgreSQL – implementacja baz danych”
- Dokumentacja PostgreSQL: <https://www.postgresql.org/docs/10/static/>